



Query Match Similarity 63.3%; Score 1017; DB 2; Length 223;  
 Best Local Similarity 90.0%; Pred: No. 1.2e-45; Matches 194; Conservative 7; Mismatches 13; Indels 0; Gaps 0;  
 Matches 194; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 98 1DLSDISCVPUVAKEVLDICGQVDLJINASVKVCPANKSLEDKKINDANFQPII 157  
 DB 10 VDLSPDSCVQVAKEVLDICGQVDLJINASVKVCPANKSLEDKKINDANFQPII 157  
 QY 158 TKALLEPMISHRTGQTVLNUINQGKQGIPPFRTYAAKHALGFFDCIRABVEEVDDVIS 217  
 DB 70 TROLLPMISRTGQTVLNUINQGKQGIPPFRTYAAKHALGFFDCIRABVEEVDDVIS 129  
 QY 218 TSPPTFIRSYHVEPEQGNWESIWKFRKTYGVHVEVEEVMTYRKQEVFMANP 277  
 DB 130 TSPPTFIRSYHVEPEQGNWESIWKFRKTYGVHVEVEEVMTYRKQEVFMANP 189  
 QY 278 1PKAAVFTVTPPEFPAVAVGKVKLNUBEG 311  
 DB 190 VPKAAVFIRTPPEFPAVAVGKVKLNUBEG 223

RESULT 2  
 Q6UXK9 PRELIMINARY; PRT; 310 AA.  
 ID Q6UXK9  
 AC  
 DT 05-JUL-2004 (TREMBLel. 27, Created)  
 DT 05-JUL-2004 (TREMBLel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLel. 27, Last annotation update)  
 DB Q6UXK9  
 OS ORFNAMESADINQ212;  
 OC Homo sapiens (Human).  
 OC Bokaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22887956; PubMed=12975309; DOI=10.1101/gr.129303;  
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Bush J.,  
 RA Chen J., Chow B., Choi C., Crowley C., Curreli B., Deneil B., Dowd P.,  
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,  
 RA Hwang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 RA Lewis L., Liao D., Robbie E., Sanchez C., Schoenfeld J.,  
 RA Sespagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,  
 RA Vandelin R., Watanabe C., Wieand D., Woods K., Xie M.H., Yansura D.,  
 RA Yu S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
 RA Godowski P.,  
 RT "The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment.";  
 RT Genome Res. 13:2265-2270 (2003).  
 CC -!- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases (SDR) family.  
 DR EMBL: AK55498; RAO8862.1; -- activity: TEA.  
 DR GO: GO:0016491; F:oxidoreductase activity; TEA.  
 DR GO: GO:0008152; P:metabolism; IEA.  
 DR InterPro: IPR002198; ADH\_short.  
 DR InterPro: IPR002347; Adh\_short\_C2.  
 DR Pfam: PF00106; adh\_short\_1.  
 DR PRINTS; PR00081; SDR001.  
 DR PROSITE; PS00061; ADH\_SHORT; UNKNOWN\_1.  
 DR SEQUENCE 310 AA; 33524 MW; 5BE703478BE20BD7 CRC64;  
 SQ

Query Match Similarity 40.5%; Score 651.5; DB 2; Length 310;  
 Best Local Similarity 43.5%; Pred: No. 1.2e-45; Matches 138; Conservative 59; Mismatches 89; Indels 31; Gaps 6;  
 Matches 138; Conservative 59; Mismatches 89; Indels 31; Gaps 6;

QY 3 1WMLMLPLLI-LGISGLLFIYQEVSRLMSKSAVQKVTITDAISLGKECARVFTG 60  
 DB 4 ITSTATPLPLRGCLGTVFGLLQWTR--GKAVIARNAVVWITGATSGLGKECARVFTG 60  
 DR EMBL: BC04126; AAH04126.1; --  
 DR EMBL: BC009679; AAH09679.1; --  
 DR HSSP; Q9ZFY9; 1FOH.

Gencore version 5.1.6  
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 16:50:09 ; Search time 41 Seconds  
 (without alignments)

Title: US-10-664-506a-5  
 Perfect score: 1607  
 Sequence: 1 MGVMAMMLMLPULLLGGISGLL.....FFFAVAVACGVKEKLNVPEEG 311

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5.

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : PIR 79;\*

1: pir1;\*

2: pir2;\*

3: pir3;\*

4: pir4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No. Score % Query Length DB ID

#### Description

hypothetical prote

hypothetical prote

hypothetical prote

probable short-chain oxidoreductase homolog

Query Match 39.9%; Score 640.5; DB 2; Length 325; Best Local Similarity 44.2%; Pred. No. 1.5e-45; Mismatches 136; Conservative 58; Mismatches 101; Indels 13; Gaps 5; Matches 136; Description

QY

3 VVAMMLMLPULL-LGSGGLIPIQYEVSRLLWSKSAVONKVVWITDAISGLGKCAEVRTTG 60

QY

19 IIRSTALPLRFLSCIGVFLFRLLQWR--GKAYLRNAAVWVITGATSGLGKCAVKYAA 75

QY

61 GARLVIAGCAWENLJYDAL-LSVADPSKTFPKULDSDISCPFDVAKVLDYGC 119

QY

76 GAKLVIAGCAWENLJYDAL-LSVADPSKTFPKULDSDISCPFDVAKVLDYGC 135

QY

120 VDILLINNSVKYKGPAHKISLEDKKTMIDNYFGPTITKALLPNMSSRTSQIVLVLNNI 179

QY

136 VDILVNNAGISRTGTTIMDTIVDVKRMTNTFGPVALKPSMVKRQHIVAISSI 195

QY

180 QKKGIGPRRTYASKKALGFDCLVAVEVYDVWVISTVSPFIR--SYVYYPEQNW 236

QY

196 QGKNSIPPSAYAZPKHATQAFFDCLRAKEMOEIEVTVISPGYIHTNLVIAITADG- 253

QY

237 EASIIKWFPRKLTIGVHEVVEBMRTRRKQEVUMANRPKAANVVRTERPEPFV 296

QY

254 -SRGGMDDTQAQGRSPVEVAQDVLAVGKCKDVLADLPLSLAVYLRTLAPGLFSL 311

QY

297 VACGVKEK 304

QY

312 MRSARKE 319

DB

oxidoreductase ypa  
 probable 3-oxacyl  
 hypothetical prote  
 3-oxoacyl-(acyl) ca  
 hypothetical prote  
 ribitol 2-dehydro  
 conserved hypothet  
 oxidoreductase of  
 hypothetical prote  
 short-chain type d  
 oxidoreductase hom  
 internalin B - Lis

#### ALIGNMENTS

RESULT 2  
 T34380  
 CSpecies: Caenorhabditis elegans  
 CDate: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T34380  
 R;Du, Z.  
 A;Description: The sequence of *C. elegans* cosmid T25G12.  
 A;Reference number: Z21515  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-537 <DUZ>  
 A;Cross-references: UNIPROT:Q22787; EMBL:U432283; PIDN: AAC69018.1; GSPDB: GN00028; CESP:T2  
 A;Experimental source: strain Bristol N2; clone T25G12  
 C;Genetics:  
 A;Gene: CESP:T25G12.2  
 A;Map position: X  
 A;Introns: 59/2; 89/1; 136/2; 279/3; 342/2; 426/3; 475/2; 526/2  
 Query Match 19.4%; Score 311.5; DB 2; Length 537;  
 Best Local Similarity 0.0%; Pred. No. 5.7e-18;  
 Matches 88; Conservative 57; Mismatches 99; Indels 49; Gaps 9;  
 QY 31 SKSAYONKVVUTDAISGIGKCARVHTGCAARLYVLUCKWNERIINYDALISVADPSK 19  
 Db 72 SKLNVYKNIKVITVAGSSGLGKSLAFELYKRGAAQVILLARSTEKLEI----CAELTKT 125  
 QY 91 F----TPKLVLDLSDISCVP--DVAKEVLDYCGVDILINNAVKVKGPAAHKTSLED 143  
 Db 126 FPLNKNKPTVYFFDITNPDKAPWAQPK----DWDVNNAGMSNRGSCQDITMP 177  
 QY 144 KKKIMDANYFGPITLTKALLPWNMSRTGQTVLUNNTOCKKGPRFTTYAASKHALGFD 203  
 Db 178 RAKAMEINLFGHVQVQIQLSLGI----SPDGCVIVTSSIQKRAPIPRGSYSAKHALQGYFD 235  
 QY 204 CIRASVEEVYVWVISTVSPTRFSYH-----VVPYQGNEHEASTIWKFPERLTYGV 252  
 Db 236 CTRASHKHLHLIV--VSAGYINTGFGSRAALITDGKVVGVDENOKK-----GY 281  
 QY 253 HPVVEAVBEMTVRKKQEWFMANIPKAIVVTRTFPPEF--FAVWVACGKEK 304  
 Db 282 SPEHSAARMISDAIRDRVSDFMAPFGARFAFLRLVFWPTLNVALYTRGKTDQ 334  
 RESULT 3  
 T34378 hypothetical protein T25G12.7 - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 29-Oct-1999 #sequence\_revision 29-oct-1999 #text\_change 09-Jul-2004  
 C;Accession: T34378  
 R;Du, Z.  
 submitted to the EMBL data library, December 1995  
 A;Description: The sequence of *C. elegans* cosmid T25G12..  
 A;Reference number: Z21515  
 A;Accession: T34378  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-262 <DUZ>  
 A;Cross-references: UNIPROT:Q22785; EMBL:U432283; PIDN: AAC69022.1; GSPDB: GN00028; CESP:T2  
 A;Experimental source: strain Bristol N2; clone T25G12  
 C;Genetics:  
 A;Gene: CESP:T25G12.7  
 A;Map position: X  
 A;Introns: 28/2; 58/1; 105/2  
 Query Match 19.1%; Score 307; DB 2; Length 262;  
 Best Local Similarity 34.8%; Pred. No. 5.4e-18;  
 Matches 80; Conservative 48; Mismatches 70; Indels 32; Gaps 8;  
 QY 10 PULLIGISGILPIKQEVSLR----WKSQAVQVKWVTDALSGIGKCARVHTGCAARLYVLUCKWNERIINYDALISVADPSK 63  
 Db 11 PALVYPLS-LYVAKVNLNRPITPGAHNLPLKUDVKVWVITGASSGIGKSLAFELYKRGAAQVILLARSTEKLEI----CAELTKT 73  
 QY 64 LVLCKVNWERLNLVYDLSWADPSKTF----TPKLVLDLSDISCVP--DVAKEVLD 116  
 Db 74 VILLARSTEKLEI----CAELTKT-----KETPPLNQNPPIVYVFFDITDSEQAPWAIEPR---- 122  
 RESULT 4  
 E83152 probable short-chain dehydrogenase PA3957 (imported) - *Pseudomonas aeruginosa* (strain I  
 C;Species: *Pseudomonas aeruginosa*  
 C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
 C;Accession: E83152  
 C;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen  
 A;Reference number: A82950; MUID:20437337; PMID:10984043  
 A;Accession: E83152  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-278 <DUZ>  
 A;Cross-references: UNIPROT:Q9H259; GB:AE004813; GB:AE004091; NID:99950134; PIDN:AAQ07;  
 A;Experimental source: strain PA01  
 C;Genetics:  
 A;Gene: PA3957  
 C;Superfamily: 1lbeta-hydroxysteroid dehydrogenase  
 Query Match 18.6%; Score 299.5; DB 2; Length 278;  
 Best Local Similarity 28.6%; Pred. No. 2.5e-17;  
 Matches 87; Conservative 57; Mismatches 99; Indels 61; Gaps 8;  
 QY 28 RUMSKSAVONKVVUTDAISGIGKCARVHTGCAARLYVLUCKWNERIINY----- 78  
 Db 4 KVSSQARYTHKVLVSGCGSGIGRALAHLRPARAGARLALIDQOALDSVQHLDHDGG 63  
 QY 79 DALFISVADPSKTFPKLVLDLSDISCVDFDAKEVLDYCGVDILINNAVKVKGPAAHKI 138  
 Db 64 EAL-----GIRCDVADADAVERAVALVERFEGGIDLVLNNAIGIHTRGTEET 110  
 QY 139 SLELDKKIMDANYFGPITLTKALLPWNMSRTGQIVLUNNTOCKKGPRFTTYAASKRA 198  
 Db 111 GLGVFRKVNANVNEFGAVICTRALPLSLERR-GQIIVVGLSITGAPLILRSAYNAASKRA 169  
 QY 199 LGFEDCIALEAVYDVKVSTVSPTFIR-----SYVTPED--GNMEEASIWKF 244  
 Db 170 HGLFDTLAMELGCGVSTLACPGFTADPLKRNALVGBGSVTRQPVQVNGSOVAS---- 224  
 QY 245 FRKLYTYGHFVVEAVBEMTVRKKQEWFMANIPKAIVVTRTFPPEFPAVWACGKEK 304  
 Db 225 -----PVEVAEATQGAARRRLVLSNVNMRALLAR-FFPRP-----EK 265  
 QY 305 LNVP 308  
 Db 266 LLVP 269  
 RESULT 5  
 AG1983 hypothetical protein all1418 [imported] - *Nostoc* sp. (strain PCC 7120)  
 C;Species: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
 A;Note: Nostoc sp. strain PCC 7120  
 C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
 C;Accession: AG1983  
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchi, T.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.; DNA Res. 8, 205-213, 2001  
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ar*  
 A;Reference number: AB1807; MUID:21595285; PMID:11759840

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: August 12, 2005, 16:51:19 ; Search time 43 Seconds

Sequence: 1 MGVMAMMLMPILLIGISGLU.....FFEAVVACGVTEKEKUNVPEEG 311

539.904 Million cell updates/sec

Title: US-10-664-506a-5  
Perfect score: 1607

Sequence: 1 MGVMAMMLMPILLIGISGLU.....FFEAVVACGVTEKEKUNVPEEG 311

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Issued\_Patents\_XA:\*

1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep: \*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep: \*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep: \*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep: \*  
5: /cgn2\_6/ptodata/1/iaa/PCITS-COMB.pep: \*  
6: /cgn2\_6/ptodata/1/iaa/backfile1.pep: \*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No. Score Query Match Length DB ID Description

RESULT 1  
US-09-634-955B-5  
; Sequence 5, Application US/09634955B  
; Patent No. 6511834  
; GENERAL INFORMATION  
; APPLICANT: Meyers, Rachel  
; APPLICANT: Cook, William James  
; TITLE OF INVENTION: 3242, 21481, 25664, 21686, NOVEL HUMAN DEHYDROGENASE  
; TITLE OF INVENTION: MOLECULES AND USES THEREFOR  
; FILE REFERENCE: M01-134  
; CURRENT APPLICATION NUMBER: US/09-634, 955B  
; CURRENT FILING DATE: 2000-08-08  
; PRIORITY APPLICATION NUMBER: 60/192, 002  
; PRIORITY FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSEQ for Windows Version 4.  
; SEQ ID NO 5  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-634-955B-5  
Query Match 100.0%; Score 1607; DB 4; Length 311;  
Best Local Similarity 100.0%; Pred. No. 8.2e-171;  
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Ov 1 MGVMAMMLMPILLIGISGLU.....FFEAVVACGVTEKEKUNVPEEG 60  
Db 1 MGVMAMMLMPILLIGISGLU.....FFEAVVACGVTEKEKUNVPEEG 60  
Ov 61 GARLVLGKNNERLENLYDALISVADPKTTPKLVILDSDISCVPPDAKEVLDCCV 120  
Db 61 GARLVLGKNNERLENLYDALISVADPKTTPKLVILDSDISCVPPDAKEVLDCCV 120  
Ov 121 DILINNAKSVKGPAHKLISLEDKKKMDANFGPITIKKALPNMISRTGQIVUANNIQ 180  
Db 121 DILINNAKSVKGPAHKLISLEDKKKMDANFGPITIKKALPNMISRTGQIVUANNIQ 180  
Ov 181 GKGPGIPRTVYASKHALGFDCLAEVB3DVWISTVSPFIRSYHVPQGNMBSI 240  
Db 181 GKGPGIPRTVYASKHALGFDCLAEVB3DVWISTVSPFIRSYHVPQGNMBSI 240  
Ov 301 VKEKUNVPEEG 311  
Db 301 VKEKUNVPEEG 311

ALIGMENTS

28 236.5 14.7 270 4 US-09-902-540-15606 Sequence 15606, A  
29 233 14.5 332 4 US-09-949-016-161 Sequence 6161, AP  
30 228.5 14.2 226 4 US-09-902-540-14378 Sequence 14378, A  
31 227 14.1 253 4 US-09-983-110-3681 Sequence 3681, AP  
32 227 14.1 284 4 US-09-07-433-4311 Sequence 4311, AP  
33 224.5 14.0 257 4 US-09-489-039A-8162 Sequence 8162, AP  
34 224 13.9 238 4 US-09-866-666-1 Sequence 1, APPL  
35 224 13.9 248 3 US-09-385-028-11 Sequence 11, APPL  
36 224 13.9 248 4 US-09-726-614-11 Sequence 11, APPL  
37 224 13.9 248 4 US-09-385-040-11 Sequence 5219, AP  
38 218.5 13.6 249 4 US-09-134-000C-5219 Sequence 4092, AP  
39 218 13.6 272 4 US-09-107-433-4092 Sequence 12822, A  
40 215 13.4 271 4 US-09-689-039A-12822 Sequence 12880, A  
41 215 13.4 292 4 US-09-489-039A-12880 Sequence 43794, A  
42 212 13.2 361 4 US-02-270-767-43794 Sequence 4397, AP  
43 209 13.0 231 3 US-09-134-001C-4337 Sequence 37, APP  
44 209 13.0 300 4 US-09-907-74A-159 Sequence 37, APP  
45 209 13.0 300 4 US-09-866-028-37 Sequence 37, APP

RESULT 2  
 US-09-816-760-5  
 ; Sequence 5, Application US/09816760  
 ; Patent No. 613555  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Meyers, Rachel  
 ; APPLICANT: Cook, William James  
 ; APPLICANT: Williamson, Mark  
 ; APPLICANT: Rudolph-Owen, Laura A.  
 ; TITLE OF INVENTION: 32142, 21481, 25964, 21686, NOVEL DEHYDROGENASE  
 ; TITLE OF INVENTION: MOLECULES AND USES THEREFOR  
 ; FILE REFERENCE: MNI-134CP  
 ; CURRENT APPLICATION NUMBER: US/09/816,760  
 ; CURRENT FILING DATE: 2001-03-23  
 ; PRIOR APPLICATION NUMBER: 09/634,955  
 ; PRIOR FILING DATE: 2000-08-08  
 ; PRIOR APPLICATION NUMBER: 60/192,002  
 ; PRIOR FILING DATE: 2000-03-24  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 311  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-816-760-5  
 ;  
 ; Query Match 100.0%; Score 1607; DB 4; Length 311;  
 ; Best Local Similarity 100.0%; Pred. No. 8.2e-11; Mismatches 0; Indels 0; Gaps 0;  
 ; Matches 311; Conservative 0; MisMatches 0; Index 0;  
 ;  
 Qy 1 MGVMAMMLPILLGIGSGLIPIQEVSLWSKSAVONKVVITDAISGIGKECARVFTG 60  
 Db 1 MGVMAMMLPILLGIGSGLIPIQEVSLWSKSAVONKVVITDAISGIGKECARVFTG 60  
 Qy 61 GARLVIGKGNWERLENLYDALISVADPSKTPKVLIDSDISCPVDAKEVLDYCGV 120  
 Db 61 GARLVIGKGNWERLENLYDALISVADPSKTPKVLIDSDISCPVDAKEVLDYCGV 120  
 Qy 121 DILINNAVKVKGPARKSISLEDKKKMDANYFGPILTAKLPMISRTGQIVLNNIQ 180  
 Db 121 DILINNAVKVKGPARKSISLEDKKKMDANYFGPILTAKLPMISRTGQIVLNNIQ 180  
 Qy 181 GKFGGIPPRTYAASKHAGFFDCITRAVEEYDVISTWSPTFTRSYVTYPEQNWEST 240  
 Db 181 GKFGGIPPRTYAASKHAGFFDCITRAVEEYDVISTWSPTFTRSYVTYPEQNWEST 240  
 Qy 241 WKEFPRKLTYGVPHPEVAAEMTRKKEVEMANPIKAATYVTRTPPEFFAVACG 300  
 Db 241 WKEFPRKLTYGVPHPEVAAEMTRKKEVEMANPIKAATYVTRTPPEFFAVACG 300  
 Qy 301 VKEKLNVPBEG 311  
 Db 301 VKEKLNVPBEG 311  
 ;  
 ; RESULT 3  
 ; US-09-838-561-5  
 ; Sequence 5, Application US/09838561  
 ; Patent No. 6627423  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Meyers, Rachel  
 ; APPLICANT: Cook, William James  
 ; APPLICANT: Williamson, Mark  
 ; APPLICANT: Rudolph-Owen, Laura A.  
 ; APPLICANT: Gimeno, Ruth  
 ; TITLE OF INVENTION: 32142, 21481, 25964, 21686, NOVEL DEHYDROGENASE  
 ; TITLE OF INVENTION: MOLECULES AND USES THEREFOR  
 ; FILE REFERENCE: MNI-134CP2  
 ; CURRENT APPLICATION NUMBER: US/09/838,561  
 ; CURRENT FILING DATE: 2001-04-18  
 ; PRIOR APPLICATION NUMBER: 09/816,760  
 ; PRIOR FILING DATE: 2001-03-23  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 15  
 ; LENGTH: 311  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-09-838-561-5  
 ;  
 ; Query Match 91.3%; Score 1467; DB 4; Length 311;  
 ; Best Local Similarity 91.3%; Pred. No. 3.6e-15;  
 ; Matches 284; Conservative 8; Mismatches 19; Indels 0; Gaps 0;

Copyright (c) 1993 - 2005 Compugen Ltd.	GenCore version 5.1.6										
Run on:	August 12, 2005, 16:41:03 ; Search time 165 Seconds (without alignments)										
Scoring table:	BLOSUM62 Gapop 10.0 , Gapext 0.5										
Searched:	2105692 seqs, 38676031 residues										
Total number of hits satisfying chosen parameters:	2105692										
Minimum DB seq length:	0										
Maximum DB seq length:	2000000000										
Post-processing:	Minimum Match 0% Maximum Match 100% Listing first 45 summaries										
Database :	A_Geneseq_16Dec04;*										
1: geneseq1980s;*	AAB47593										
2: geneseq2190s;*	AAB47593 standard; protein; 311 AA.										
3: geneseq2000s;*	XX										
4: geneseq2001s;*	XX										
5: geneseq2002s;*	AC										
6: geneseq2003sab;*	AAB47593;										
7: geneseq2004sab;*	XX										
8: geneseq2004s;*	XX										
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	SUMMARIES										
Result No.	Score	Query	Length	DB	ID	Description	ALIGNMENTS				
1	1607	100.0	311	4	AAB47593	RESULT 1	AAB47593	XX	XX	XX	XX
2	1607	100.0	311	6	ABP8046	ID	AAB47593 standard; protein; 311 AA.	XX	XX	XX	XX
3	1607	100.0	311	8	ADG0199	XX	XX	XX	XX	XX	XX
4	1601	99.6	311	3	ABP80911	XX	XX	XX	XX	XX	XX
5	1601	99.6	345	3	AAB08948	XX	XX	XX	XX	XX	XX
6	1467	91.3	311	6	ABP8047	XX	XX	XX	XX	XX	XX
7	1467	91.3	311	8	ADG0201	XX	XX	XX	XX	XX	XX
8	1258	246	8	ADM3425	XX	XX	XX	XX	XX	XX	XX
9	651.5	40.5	310	3	AAB37650	W0200172976-A2.	XX	XX	XX	XX	XX
10	651.5	40.5	310	3	ADCT8473	XX	XX	XX	XX	XX	XX
11	651.5	40.5	310	4	AAB08238	XX	XX	XX	XX	XX	XX
12	651.5	40.5	310	5	ABP84835	XX	XX	XX	XX	XX	XX
13	651.5	40.5	310	5	ABP5441	XX	XX	XX	XX	XX	XX
14	651.5	40.5	310	6	ABU1616	XX	XX	XX	XX	XX	XX
15	651.5	40.5	310	6	ABU71471	XX	XX	XX	XX	XX	XX
16	651.5	40.5	310	6	ABU71917	XX	XX	XX	XX	XX	XX
17	651.5	40.5	310	6	ABO01800	XX	XX	XX	XX	XX	XX
18	651.5	40.5	310	6	ABU54373	XX	XX	XX	XX	XX	XX
19	651.5	40.5	310	6	ABO47388	XX	XX	XX	XX	XX	XX
20	651.5	40.5	310	6	ABU6525	XX	XX	XX	XX	XX	XX
21	651.5	40.5	310	6	ABU67371	XX	XX	XX	XX	XX	XX
22	651.5	40.5	310	6	ABO14891	XX	XX	XX	XX	XX	XX
23	651.5	40.5	310	6	ABU68648	XX	XX	XX	XX	XX	XX
24	651.5	40.5	310	6	ABO24830	XX	XX	XX	XX	XX	XX
25	651.5	40.5	310	6	ABO29358	XX	XX	XX	XX	XX	XX

Ada18214 Human sec  
Abc32782 Human sec  
Abc34842 Human PRO  
Ada16189 Human sec  
Ada2334 Human sec  
Abc1720 Human PRO  
Ada16613 Human sec  
Ada13042 Human sec  
Ada41910 Human sec  
Ada17257 Human sec  
Ada42760 Human sec  
Abc17581 Human PRO  
Abt7679 Human sec  
Abt7815 Human sec  
Abc28461 Human sec  
Ada39661 Human sec  
Adc40175 Human sec  
Adc19003 Human sec  
Adc34299 Human sec  
Adc29354 Human sec

Abc47593 Human sec  
Abt77650 Human sec  
Abc1917 Human sec  
Abc3425 Human sec  
Abc28461 Human sec  
Ada39661 Human sec  
Adc40175 Human sec  
Adc19003 Human sec  
Adc34299 Human sec  
Adc29354 Human sec

New human dehydrogenase molecules designated DHDR-1, DHDR-2, DHDR-3 and DHDR-4 for finding treatment for dehydrogenase-associated disorders such as viral and cellular proliferation disorders.

CC The sequences given in AAH43560-67 encode human dehydrogenase (DHDR)-1-4.  
 CC Modulators of DHDR activity are used to treat a viral or cellular  
 CC proliferation disorder. The DHDR proteins can also be used to treat  
 CC disorders characterized by over or under production of DHDR proteins, for  
 CC example dehydrogenase-associated disorders such as CNS, cardiac,  
 CC muscular, cellular growth, differentiation or migration, neurological,  
 CC immune, humoral or viral disorders. The molecules can be used in  
 XX pharmacogenetics

SQ Sequence 311 AA:  
 Query Match 100.0%; Score 1607; DB 4; Length 311;  
 Best Local Similarity 100.0%; Pred. No. 7.4e-163; Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1 MGVMAMMLMLPLLIGLISGLFLYQEVSRILWSKSAVONKVVITDAISGGKGECARVFTG 60  
 QY 61 GARLVLCKGKWRLENLYDALISVADPSKTFPKVLIDLSISCVDPVAKEVLDCCGV 120  
 Db 61 GARLVLCKGKWRLENLYDALISVADPSKTFPKVLIDLSISCVDPVAKEVLDCCGV 120  
 QY 121 DILINNASVKKGPAHKISLEDKKIMDANYGFPITLKLPMNSRRTGQIVLNHQ 180  
 Db 121 DILINNASVKKGPAHKISLEDKKIMDANYGFPITLKLPMNSRRTGQIVLNHQ 180  
 QY 181 GKFGGIPPFRTTYAASKHAGFFDCIRAEVEEDVYDVTWVSPTPRSYHYPYPOQGNWEASI 240  
 Db 181 GKFGGIPPFRTTYAASKHAGFFDCIRAEVEEDVYDVTWVSPTPRSYHYPYPOQGNWEASI 240  
 QY 241 WKFFRKLTGTVHPVVEAEVWMTVRKQEVFMANPIKPAVYVRFPEFFFAVWAGC 300  
 Db 241 WKFFRKLTGTVHPVVEAEVWMTVRKQEVFMANPIKPAVYVRFPEFFFAVWAGC 300  
 QY 301 VKEKLNVEEG 311  
 Db 301 VKEKLNVEEG 311

RESULT<sup>2</sup>

ID ABP58045

ID ABP58046

AC

XX

DT

XX

DE

XX

KW

KW

KW

KW

OS

XX

PN

XX

PD

XX

PF

XX

PR

XX

PA

XX

PT

XX

DR

XX

PT

PT obesity or diabetes, comprises assaying the ability of the compound to  
 PT modulate dehydrogenase-2 (DHDR-2) nucleic acid expression or DHDR-2  
 PT polypeptide activity.

CC Claim 13; Fig 3; 73pp; English.

CC The present sequence is the protein sequence of human dehydrogenase-2  
 CC (DHDR-2). The invention is based on the discovery that expression of the  
 CC DHDR-2 gene is downregulated during cold exposure, i.e. during  
 CC thermogenesis, and is also downregulated in cells expressing thermogenic  
 CC coactivator PGC-1. DHDR-2 molecules, as part of the thermogenic  
 CC signalling pathway, modulate thermogenesis and are useful as targets and  
 CC therapeutic agents for the modulation of thermogenesis, e.g. expenditure  
 CC of energy and the treatment of body weight disorders. A claimed method  
 CC for identifying a compound capable of treating a body weight disorder  
 CC comprises assaying the ability of the compound to modulate DHDR-2 nucleic  
 CC acid expression or DHDR-2 polypeptide activity, e.g. by detecting  
 CC mitochondrial activity of a cell or detecting thermogenesis in a cell.  
 CC The body weight disorder is obesity, overweight, diabetes, insulin  
 CC resistance, cachexia or anorexia. A claimed method for identifying a cell  
 CC compound capable of modulating thermogenesis involves contacting a cell  
 CC which expresses DHDR-2, especially a primary muscle cell, a C2C12 myocyte  
 CC and a C2C12 myotube with a test compound and assaying the ability of the  
 CC compound to modulate DHDR-2 expression or activity. The modulator is a  
 CC DHDR-2 polypeptide, its fragment or variant, an antisense DHDR-2 nucleic  
 CC acid, a ribozyme or a DHDR-2 nucleic acid. The DHDR-2 modulator is used  
 CC in a claimed method of creating a body weight disorder, and may be  
 CC administered using a gene therapy vector. The methods and DHDR-2  
 CC modulators are useful for diagnosing, preventing and treating a body  
 CC weight disorder or for modulating thermogenesis

XX Sequence 311 AA;

SQ

Query Match

100.0%;

Score 1607;

DB 6;

Length 311;

Best Local Similarity

100.0%;

Pred. No. 7.4e-163;

Matches 311;

Conservative 0;

Mismatches 0;

Indels 0;

Gaps 0;

Db

1 MGVMAMMLMLPLLIGLISGLFLYQEVSRILWSKSAVONKVVITDAISGGKGECARVFTG 60  
 QY 61 GARLVLCKGKWRLENLYDALISVADPSKTFPKVLIDLSISCVDPVAKEVLDCCGV 120  
 Db 61 GARLVLCKGKWRLENLYDALISVADPSKTFPKVLIDLSISCVDPVAKEVLDCCGV 120  
 QY 121 DILINNASVKKGPAHKISLEDKKIMDANYGFPITLKLPMNSRRTGQIVLNHQ 180  
 Db 121 DILINNASVKKGPAHKISLEDKKIMDANYGFPITLKLPMNSRRTGQIVLNHQ 180  
 QY 181 GKFGGIPPFRTTYAASKHAGFFDCIRAEVEEDVYDVTWVSPTPRSYHYPYPOQGNWEASI 240  
 Db 181 GKFGGIPPFRTTYAASKHAGFFDCIRAEVEEDVYDVTWVSPTPRSYHYPYPOQGNWEASI 240  
 QY 241 WKFFRKLTGTVHPVVEAEVWMTVRKQEVFMANPIKPAVYVRFPEFFFAVWAGC 300  
 Db 241 WKFFRKLTGTVHPVVEAEVWMTVRKQEVFMANPIKPAVYVRFPEFFFAVWAGC 300  
 QY 301 VKEKLNVEEG 311  
 Db 301 VKEKLNVEEG 311
RESULT<sup>3</sup>

ID ADG40199

ID ADG40199

AC

XX

DT

XX

DE

XX

KW

KW

KW

KW

OS

XX

PN

XX

PD

XX

PF

XX

PR

XX

PA

XX

PT

XX

DR

XX

PT

XX

CC identifying a compound useful for treating a bodyweight disorder, e.g.  
 CC

-continued

aac aac atc caa gcg aag ttt gga atc ccg ttc cgc aca gct tat gca	576
Asn Asn Ile Gln Ala Lys Phe Gly Ile Pro Phe Arg Thr Ala Tyr Ala	
180 185 190	
gcc tct aag cat gcc gtc atg ggc ttc ttt gac tgc ctc cga gcc gag	624
Ala Ser Lys His Ala Val Met Gly Phe Phe Asp Cys Leu Arg Ala Glu	
195 200 205	
gtt gag gaa tac gat gtt gtg gtc agc acc gtg agc cca act ttc atc	672
Val Glu Glu Tyr Asp Val Val Val Ser Thr Val Ser Pro Thr Phe Ile	
210 215 220	
cgc tcc tac cgt gct tcc cct gag caa aga aac tgg gag aca tcc att	720
Arg Ser Tyr Arg Ala Ser Pro Glu Gln Arg Asn Trp Glu Thr Ser Ile	
225 230 235 240	
tgt aaa ttc ttc tgc agg aag cta gcc tat ggc gtg cac ccg gtg gag	768
Cys Lys Phe Phe Cys Arg Lys Leu Ala Tyr Gly Val His Pro Val Glu	
245 250 255	
gtg gct gag gaa gtg atg cgc aca gta cgg agg aag aag caa gag gtg	816
Val Ala Glu Val Met Arg Thr Val Arg Arg Lys Lys Gln Glu Val	
260 265 270	
tcc atg gcc aac ccg gtt cct aag gct gcc gtg ttc atc cgc acc ttc	864
Phe Met Ala Asn Pro Val Pro Lys Ala Ala Val Phe Ile Arg Thr Phe	
275 280 285	
tcc cct gag ttc ttc gct gtg gtg gcc tgt ggg gtg aag gag aag	912
Phe Pro Glu Phe Phe Ala Val Val Ala Cys Gly Val Lys Glu Lys	
290 295 300	
ctc aat gtc cca gaa gag ggt	933
Leu Asn Val Pro Glu Glu Gly	
305 310	

*Alt + DMS**Refield*

## What is claimed:

1. An isolated nucleic acid molecule comprising the 35 nucleotide sequence set forth in SEQ ID NO:4, or a full complement thereof.
2. An isolated nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:6, or a full complement thereof.
3. An isolated nucleic acid molecule consisting of the nucleotide sequence set forth in SEQ ID NO:4, or a full complement thereof.
4. An isolated nucleic acid molecule consisting of the nucleotide sequence set forth in SEQ ID NO:6, or a full complement thereof. 45
5. An isolated nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:5, or a full complement thereof.
6. An isolated nucleic acid molecule which encodes a 50 polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:5, or a full complement thereof.
7. An isolated nucleic acid molecule comprising the nucleotide sequence contained in the plasmid deposited with ATCC® as Accession Number PTA-1845, or a full complement thereof. 55
8. An isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:4 or SEQ ID NO:6, or a full complement thereof, wherein said nucleic acid molecule encodes a polypeptide having a dehydrogenase activity. 60

9. An isolated nucleic acid molecule consisting of a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:4 or SEQ ID NO:6, or a full complement thereof, wherein said nucleic acid molecule encodes a polypeptide having a dehydrogenase activity.
10. An isolated nucleic acid molecule encoding a polypeptide comprising an amino acid sequence which is at least 95% identical to the amino acid sequence of SEQ ID NO:5, or a full complement thereof, wherein said polypeptide has a dehydrogenase activity.
11. An isolated nucleic acid molecule encoding a polypeptide consisting of an amino acid sequence which is at least 95% identical to the amino acid sequence of SEQ ID NO:5, or a full complement thereof, wherein said polypeptide has a dehydrogenase activity.
12. An isolated nucleic acid molecule comprising the nucleic acid molecule of any one of claims 1-4, and a nucleotide sequence encoding a heterologous polypeptide.
13. A vector comprising the nucleic acid molecule of any one of claims 1-4.
14. The vector of claim 13, which is an expression vector.
15. A host cell transfected with the expression vector of claim 14.
16. A method of producing a polypeptide having dehydrogenase activity comprising culturing the host cell of claim 15 under conditions in which the nucleic acid molecule is expressed, thereby expressing the polypeptide.

\* \* \* \* \*

-continued

1	5	10	15
Gln	Asn	Gln	Val
20		Ala	Val
		Val	Thr
		Gly	Gly
		Ala	Thr
		Gly	Ile
		Ile	Gly
			Lys
Ala	Ile	Ser	Arg
35		Glu	Glu
		Leu	Leu
		His	Leu
		Gly	Cys
		Asn	Val
		Val	Ile
			Ala
Ser	Arg	Lys	Leu
50		Asp	Arg
		Leu	Thr
		Ala	Ala
		Val	Asp
		Glu	Leu
			Arg
Ser	Gln	Pro	Pro
65		Ser	Ser
		Thr	Gln
		Val	Thr
		Ala	Ile
		Gln	Cys
		Asn	
Ile	Arg	Lys	Glu
85		Glu	Glu
		Val	Asn
		Asn	Leu
		Val	Val
		Lys	Ser
		Thr	Leu
		Ala	
Lys	Tyr	Gly	Ile
100		Asn	Asn
		Phe	Phe
		Leu	Val
		Asn	Asn
		Ala	Gly
		Gly	Gln
		Gln	Phe
Met	Ala	Pro	Ala
115		Glu	Glu
		Asp	Ile
		Thr	Ala
		Lys	Gly
		Trp	Gln
		Gln	Ala
		Ala	Val
		Ile	
Glu	Thr	Asn	Leu
130		Thr	Gly
		Phe	Thr
		Tyr	Met
		Cys	Lys
		Ala	Val
		Tyr	Asn
Ser	Trp	Met	Lys
145		Asp	His
		Gly	Gly
		Ser	Ile
		Val	Asn
		Ile	Ile
		Val	Leu
Leu	Asn	Asn	Gly
165		Phe	Pro
		Thr	Ala
		Ala	His
		Ser	Gly
		Ala	Ala
		Arg	Ala
Gly	Val	Tyr	Asn
180		Leu	Thr
		Lys	Thr
		Met	Ala
		Leu	Thr
		Trp	Ala
		Ser	Ser
Gly	Val	Arg	Ile
195		Asn	Cys
		Val	Ala
		Pro	Gly
		Thr	Ile
		Tyr	Ser
		Gln	Thr
Ala	Val	Asp	Asn
210		Tyr	Gly
		Glu	Leu
		Gly	Gln
		Thr	Met
		Phe	Glu
		Met	Ala
Phe	Glu	Asn	Ile
225		Pro	Ala
		Lys	Arg
		Val	Gly
		Leu	Pro
		Pro	Glu
		Glu	Ile
Pro	Leu	Val	Cys
245		Phe	Leu
		Leu	Ser
		Pro	Ala
		Ala	Ser
		Phe	Ile
		Thr	Gly
Gln	Leu	Ile	Asn
260		Val	Asp
		Gly	Gly
		Gln	Ala
		Leu	Tyr
		Thr	Arg
		Asn	Phe
Thr	Ile	Pro	Asp
275		His	Asp
		Asn	Trp
		Pro	Val
		Gly	Ala
		Gly	Aep
		Ser	Ser
Phe	Ile	Lys	Lys
290		Val	Lys
		Glu	Ser
		Leu	Lys
		Lys	Gln
		Gln	Ala
		Arg	Leu

What is claimed:

1. A method for identifying a compound which binds to a polypeptide comprising the amino acid sequence of SEQ ID NO:2;

the method comprising:

- contacting the polypeptide, or a cell expressing the polypeptide with a test compound under conditions suitable for binding; and
- detecting binding of the test compound to the polypeptide.

2. A method for identifying a compound which binds to a polypeptide consisting of the amino acid sequence of SEQ ID NO:2;

the method comprising:

- contacting the polypeptide, or a cell expressing the polypeptide with a test compound under conditions suitable for binding; and
- detecting binding of the test compound to the polypeptide.

50 3. A method for identifying a compound which binds to a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:10 or 12, and wherein said polypeptide has a dehydrogenase activity;

the method comprising:

- contacting the polypeptide, or a cell expressing the polypeptide with a test compound under conditions suitable for binding; and
- detecting binding of the test compound to the polypeptide.

4. A method for identifying a compound which binds to a polypeptide comprising an amino acid sequence which is at least 95% identical to the entire length of the amino acid sequence of SEQ ID NO:11, and wherein said polypeptide has a dehydrogenase activity;

the method comprising:

- i) contacting the polypeptide, or a cell expressing the polypeptide with a test compound under conditions suitable for binding; and
- ii) detecting binding of the test compound to the polypeptide.

5. A method for identifying a compound which binds to a polypeptide encoded by the nucleotide sequence contained in the plasmid deposited with ATCC® as Accession Number PTA-3216;

the method comprising:

- a) contacting the polypeptide, or a cell expressing the polypeptide with a test compound under conditions suitable for binding; and
- b) detecting binding of the test compound to the 15 polypeptide.

6. The method of any one of claims 1, 2, 3, 4, and 5, wherein said detection is by direct binding.

7. The method of claim 6, wherein said direct binding is determined by an immunoprecipitation.

8. The method of claim 6, wherein said direct binding is determined by a yeast two-hybrid assay.

9. The method of any one of claims 1, 2, 3, 4, or 5, wherein said detection is by the use of a competition binding assay.

10. The method of any one of claims 1, 2, 3, 4, or 5, wherein said cell is a tumor cell.

11. The method of any one of claims 1, 2, 3, 4, or 5,  
10 wherein said detection is by the use of an assay for an  
activity of the polypeptide consisting of the amino acid  
sequence of SEQ ID NO:11.

12. The method of claim 11, wherein said activity is cellular growth or proliferation.

13. The method of claim 11, wherein said activity is cellular signaling.

14. The method of claim 11, wherein said activity is modulation of viral gene expression.

\* \* \* \* \*

-continued

---

35	40	45
Ser Leu		
50		
<210> SEQ ID NO 34		
<211> LENGTH: 57		
<212> TYPE: PRT		
<213> ORGANISM: Artificial Sequence		
<220> FEATURE:		
<223> OTHER INFORMATION: hypothetical protein domain		
<400> SEQUENCE: 34		
Phe Tyr Lys Pro Asn Leu Glu Gln Tyr Gln His Arg Trp Thr Val Val		
1	5	10
Ser Gly Gly Thr Asp Gly Ile Gly Lys Ala Tyr Thr Leu Glu Leu Ala		
20	25	30
Lys Arg Gly Leu Arg Lys Phe Val Leu Ile Gly Arg Asn Pro Lys Lys		
35	40	45
Leu Asp Ser Val Lys Ser Glu Ile Glu		
50	55	
<210> SEQ ID NO 35		
<211> LENGTH: 45		
<212> TYPE: PRT		
<213> ORGANISM: Artificial Sequence		
<220> FEATURE:		
<223> OTHER INFORMATION: glucose-1-dehydrogenase domain		
<400> SEQUENCE: 35		
Thr Leu Glu Met Ile Pro Ala Lys Glu Ile Gly Phe Ala Asp Gln Val		
1	5	10
Ala Asn Val Ala Arg Phe Leu Cys Ser Asp Leu Ala Asp Tyr Ile His		
20	25	30
Gly Thr Thr Ile Tyr Val Asp Gly Gly Met Thr Asn Tyr		
35	40	45

---

## What is claimed:

1. An isolated nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:10, or a complement thereof.
2. An isolated nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:11, or a complement thereof.
3. An isolated nucleic acid molecule comprising the nucleotide sequence contained in the plasmid deposited with ATCC as Accession Number PTA-3216.
4. An isolated nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:12, or a complement thereof.
5. An isolated nucleic acid molecule consisting of the nucleotide sequence set forth in SEQ ID NO:10, or a complement thereof.
6. An isolated nucleic acid molecule consisting of the nucleotide sequence set forth in SEQ ID NO:12, or a complement thereof.
7. An isolated nucleic acid molecule which encodes a polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:11, or a complement thereof.
8. An isolated nucleic acid molecule comprising the nucleic acid molecule of any one of claims 1, 2, 3, 4, 5, 6, or 7, and a nucleotide sequence encoding a heterologous polypeptide.

9. A vector comprising the nucleic acid molecule of any one of claims 1, 2, 3, 4, 5, 6, or 7.

10. The vector of claim 9, which is an expression vector.

11. A host cell transfected with the expression vector of claim 10.

12. A method of producing a polypeptide selected from the group consisting of:

- a) a polypeptide comprising or consisting of the amino acid sequence set forth in SEQ ID NO:11;
- b) a polypeptide encoded by a nucleic acid molecule comprising or consisting of the nucleotide sequence set forth in SEQ ID NO:10 or SEQ ID NO:12; and
- c) a polypeptide encoded by a nucleic acid molecule comprising or consisting of the nucleotide sequence contained in the plasmid deposited with ATCC as Accession Number PTA-3216,

the method comprising culturing the host cell of claim 10 in an appropriate culture medium to, thereby, produce the polypeptide.

13. A kit comprising the nucleic acid molecule of any one of claims 1, 2, 3, 4, 5, 6, or 7 and instructions for use.

\* \* \* \* \*

